Editorial

The International Journal of Web Applications (IJWA) has completed one year. In the four issues, we have published a large number of papers on the different web applications. It is evident that the area ‘Web World” has been experiencing different connotations with ever-increasing applications. We hope that the IJWA would address all possible significant web applications in the coming years.

In the last issue of this year, the IJWA has published a few very crucial research themes. In the first paper, “A Visual Ontology Query Language to Bridge the Semantic Gap in Content-based Image Retrieval” Awang Iskandar has addressed the semantic gap problem in the Content-based image retrieval (CBIS). To offset the problem of syntax learning in ontology, she has employed the Visual Ontology Query Language (VOQL) that interprets a visual image query into SPARQL.

Haroon Altarawneh, Moh’d M.I. Tarawneh and Sattam Allahawiah have addressed the e-commerce adoption in the web sites in their paper on “Impact of the E-Commerce on Business Pressures in Jordanian SMEs”. Dais George and Sebastian George have discussed the omni presence of heavy tails in Web Server Data and proposed a new tail index. They proved that the proposed SmooWeighted Hill estimator has least bias and least asymptotic variance.

Rose Tinabo, Fredrick Mtenzi, Ciaran O’Driscoll and Brendan O’Shea in their paper on “Protection of Personal Identifiable Information” proposed an algorithm which will enable data to be shared among different users for different purposes such as doing research and support for public benefit purposes such as hospital treatment.

Ernesto Exposito, Myriam Lamolle and Jorge Gómez-Montalvo have addressed the interoperability issues in the automatic deployment of distributed multimedia systems. They proposed an ontology framework for distributed multimedia systems design and deployment driven by user requirements in their paper on “Introducing an Ontology Driven Architecture for Distributed Multimedia Systems engineering”.

In the recent period, Web-based mapping technologies such as google maps have been increasingly utilized for traveler information systems. Yao-Jan Wu and Yinhai Wang in their paper on “An Interactive Web-based System for Urban Traffic Data Analysis” have presented a Google-map-based online system for urban traveler information broadcast and analysis. In the last paper, Dennis Lupiana, Ciaran O’Driscoll, and Fredrick Mtenzi have proposed UbiComp taxonomy to address the functionality difficulties in Ubiquitous Computing.

We hope to publish break-through and cutting technology papers in web applications in the coming issues.

Editors

A Visual Ontology Query Language to Bridge the Semantic Gap in Content-based Image Retrieval

DNF Awang Iskandar
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan
Sarawak, Malaysia
dnfaiz@fit.unimas.my

ABSTRACT: *Interest in the production and potential of digital images has increased greatly in the past decade. The main goal of a CBIR application is to find an image or a set of images that satisfy a user's information need. This leads to the semantic gap problem, which is the difficulty of relating high level human interpretations with low-level recorded visual features. The synergy between ontology and region-based image annotations is possible to reduce the gap between image features and high-level semantics. Ontology query languages are used to retrieve information stored in ontologies. Unfortunately, users need to learn the syntax before being able to query using current ontology query languages. To overcome this problem, we present the Visual Ontology Query Language (VOQL) that interprets a visual image query into SPARQL. We show that VOQL can be used to retrieve desired images.*

Keywords: Ontology, Query language, CBIR, Semantic gap.

Received: 18 May 2009, Revised 11 June 2009, Accepted 18 June 2009

1. Introduction

Most effort to minimize the semantic gap has focused on automatic image annotation (Barnard & Forsyth 2001; Jelmini and Marchand-Maillet 2003; Troncy *et al.* 2007). The images are annotated by using keywords or described formally using an ontology (Hyvönen, Saarela, Styrman and Viljanen 2003; Schreiber *et al.* 2001; Troncy *et al.* 2007). According to Brewster *et al.* (2004), an "ontology defines concepts' properties and the relationships among the concept". Approaches to bridge the semantic gap can be top-down, bottom-up or a combination of both (Hare, Lewis, Enser and Sandom 2006; Hare, Sinclair, Lewis, Martinez, Enser and Sandom 2006). Building an ontology can be seen as addressing the semantic gap problem from the top down since ontology facilitates structured knowledge presentation. As an example, an ontology can be used to semantically annotate images in a family photo album by structuring the information according to event, venue, date and people.

An ontology is composed of objects and conceptual relations between the objects. Ontology is the theory of objects in terms of the criteria which allow one to distinguish between different types of objects and their relationships, dependencies, and properties. Ontologies encode the relational structure of concepts which one can use to describe and reason about aspects of the world. This makes them eminently suitable to many problems such as image retrieval, which require prior knowledge to be modelled and utilised in both a descriptive and prescriptive capacity. Ontology is constructed to avoid ambiguous situations, hence it needs a semantic sense in order to be meaningful. This can be achieve by specifying an established rule, principle, or law Gandon (2002)— known as axioms.

Web ontology languages have been proposed as part of research related to the Semantic Web. XML, RDF, RDF Schema, OIL and DAML+OIL are among the earliest web ontology languages, while OWL is the current W3C recommendation.[1] RDF is used to represent information and to exchange knowledge on the Web. At a higher level, OWL is used to publish and share sets of terms called ontologies, supporting advanced Web search, software agents and knowledge management. OWL is the most prominent ontology language (EscÓrcio and Cardosa 2007). Ontology query languages allow expressions to be written and can be evaluated against an ontology. The queries can be used by knowledge management applications as a basis for inference actions. Existing ontology query languages include OntoQL, SPARQL, DQL (previous version of DAML+OIL), SeRQL, TRIPLE, RDQL, N3, and Versa. The SPARQL Protocol and RDF Query Language (SPARQL) is an ontology protocol[2] and query language[3] that has been adopted by W3C as the means to access and query ontologies built using RDF. Currently, SPARQL has been extended to support OWL ontologies.

In this paper, we focus on the synergy between ontology and image annotations with the aim of reducing the gap between image features and high-level semantics. Ontologies ease information retrieval. They are used to mine, interpret, and organise knowledge. An ontology may be seen as a knowledge base that can be used to improve the image retrieval process, and conversely keywords obtained from automatic tagging of image regions may be useful for creating an ontology. We engineer an ontology that surrogates concepts derived from image feature descriptors. We test the usability of the constructed ontology by querying the ontology via the Visual Ontology Query Interface, which has a formally specified grammar known as the Visual Ontology Query Language (Awang Iskandar 2009). A detail explanation of VOQI can be found in Awang Iskandar et al. (2008b). We show that synergy between ontology and image annotations is possible and this method can reduce the gap between image features and high-level semantics by providing the relationships between objects in the image. The main significant of using VOQL is that user is able to present an image search query to an ontology database without having to learn the ontology query languages.

The remainder of this paper is organised as follows. In Section 2, we present the background on existing CBIR systems and shape descriptors that we use. In Section 3, we describe the VOQL framework used in conducting our experiments. In Section 4, we present the image ontology. In Section 5, we explain the VOQL specification. In Section 6 we present and discuss experimental results, and conclude in Section 7 with a discussion of our findings and suggestions for future work.

2. Background

Image retrieval is the field of study concerned with searching and retrieving digital images from an image corpus. This field has been explored since the 1960s (Rosenfeld 1969; Tamura and Mori 1977). Early image retrieval techniques were generally based on the textual descriptions and captions of images. Today, image retrieval is one of the demanding applications that develop along with the advancement of digital imaging technologies. The exponential number of digital images produced by consumers need to be searched and allocated for future use.

Consequently, image retrieval has been diversely researched and involves several interdisciplinary research areas such as digital image processing and analysis, computer vision, information retrieval, cognitive science, computer graphics, mathematics, education and artificial intelligence.

There are various image retrieval techniques. These techniques can be categorised according to whether they are based on text, content, multimodal fusion, or semantic concepts. We differentiate these techniques by the type of features that are used to represent the images as well as the approaches that are used to retrieve similar images. The text based image retrieval techniques use keywords, the CBIR techniques use low-level image features, the multimodal fusion techniques use a combination of various image representative features, and the semantic-based techniques use concepts.

Queries for images can be based on text descriptions or image content. Text-based description queries are posed to a text-based image retrieval system, whereas the content-based image queries are posed to a CBIR system. Text-based queries can be formulated in free-text or according to a query structure. Free-text queries are normally formulated for retrieving images using the full-text information retrieval approach. Structured queries are used in image retrieval systems that are based on a particular structured representation such as XML, RDF, OWL or a database format. Content-based image queries use the extracted image features as the query examples. These image features are compared to the image features in the database, and similar images are retrieved as answers.

[1] http://www.w3.org/TR/owl-guide
[2] http://www.w3.org/TR/rdf-sparql-protocol
[3] http://www.w3.org/TR/rdf-sparql-query

Nowadays, ontologies are used to appropriately represent a structured knowledge for a domain (Hare, Sinclair, Lewis, Martinez, Enser and Sandom 2006). There have been several ongoing research projects relating to the use of ontologies for image retrieval, adopting the idea of the semantic web. Image retrieval using an ontology is a form of structured text information retrieval. The ontology can be represented by various ontology representation languages, and XML is the base language used for constructing an ontology. The integration of an ontology in image retrieval can either be used as a guide (for example WordNet) during the retrieval process or as a repository that can be queried from (Hollink *et al.* 2003; Jiang *et al.* 2004; Wang *et al.* 2006; Harit *et al*. 2005).

Town (2006) shows that using ontologies to relate semantic descriptors to their parametric representations for visual image processing leads to an effective computational and representational mechanism. The ontology implements the hierarchical representation of the domain knowledge for a surveillance system. Pre-annotated surveillance video training data and its visual descriptors are incorporated in the ontology. The ontology is used to feed information to the Bayesian inference network for tracking movement. Town also proposed an ontological query language, namely OQUEL. The query approach using OQUEL is similar to the approach presented by HyvÖnen, Saarela and Viljanen (2003) who implemented a web-based system to retrieve the images using an ontology—known as Ontagator. Image query is done using a view-based search followed by image recommendations. In the search process, users view the ontology and select the class of interest. The system will return all images related to the class. After finding a class of interest, the semantic ontology model together with image instance data are used to discover the relations between a selected image and other images in the repository. These images are then presented to the user.

Liu et al. (2004) also implemented a web-based system to retrieve the images with an ontology. Search for the matching images is done by processing a text-based query. The ontology query engine is written in RDF Data Query Language (RQDL) provided by the Jena Toolkit.[4]

3. VOQL Framework

The overall framework for VOQL is depicted in Figure 1. It consists of six components that are based on a CBIR system and ontology engineering. The CBIR system components are: an *image collection*, a *feature extraction component*, a *semantic interpretation component* and a *user interface*. The ontology engineering component includes an *ontology construction component* and a *retrieval component*. The *feature extraction component* is used to extract the shape features of the regions from the images in the collection. Based on the extracted shape features, we obtain the semantic interpretation and represent it in the form of region tags. These region tags are then processed by the *ontology construction component*. The *retrieval*

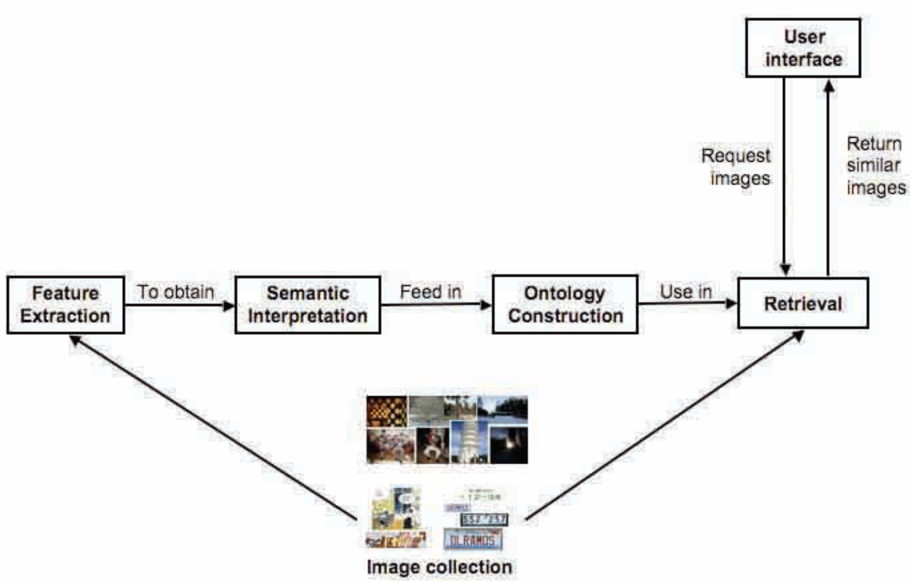

Figure 1. The VOQL Framework

[4] http://jena.sourceforge.net

component is used to process the image retrieval queries. It also serves as the means to validate the usability of the ontology. During retrieval, the constructed ontology is used as the knowledge base to answer queries formulated using the user interface. The query answers are in the form of image references. The matching images are retrieved from the image collection and displayed to the user.

4. The Image Ontology

To reduce the problem of object segmentation, we test our approach on a domain where regions are easily separated: a collection of comic strips. In this domain, objects and characters comprise multiple regions of approximately uniform colour. We have created an image collection that consists of comic strip panels from the Goats comic.[5] These include extended information that describes every panel. This description assists us in performing relevance judgements on our retrieval results. The collection consists of 452 coloured strips, each containing one to five panels. Dividing the strips into panels gives us a total of 1440 panels. We tested the retrieval effectiveness using 1115 regions extracted from 202 panels. From this point onwards, we refer to individual panels as images.

The objects in the comics have relatively consistent size and orientation, guiding our choice of the following region-based and contour-based shape features: the region area; the mean grey level value of the pixels in the region; the compactness (also known as circularity ratio or thinness ratio) of the region; and shape boundary of the region. Compactness reflects how circular the shape is. A compactness value of one indicates that the shape is a circle and zero compactness signifies a narrow shape with infinite width. Shape boundary is represented by the keypoints surrounding the shape.

We adopted the equal-weight linear combination technique from our previous work Awang Iskandar *et al*. (2008a) to recognise and label five concepts representing the main characters in the Goats comic strips Bob (an alien), Diablo (a chicken), Fineas (a fish), Jon (aperson) and Oliver (a chick). This technique was compared with classification using machine learning algorithms in combining shape features, and we found that an equal-weight linear combination of shape features is simpler and at least as effective as using amachine learning algorithm Awang Iskandar *et al*. (2008a).

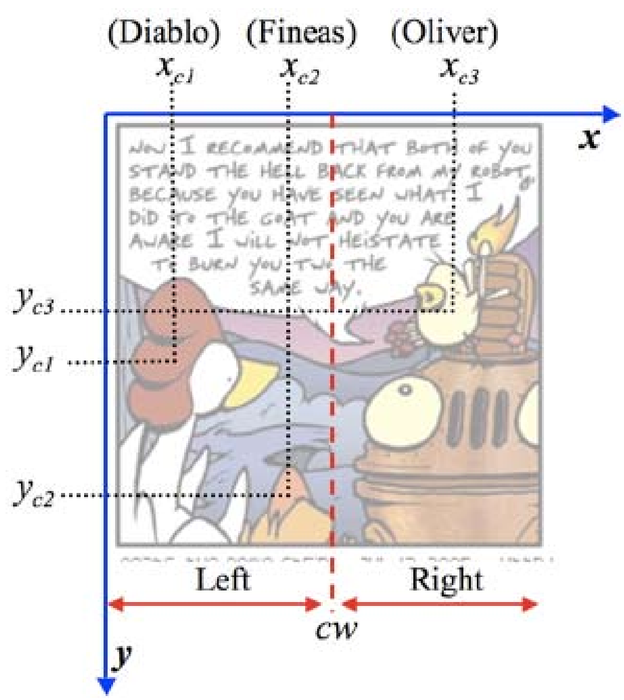

Figure 2. Oliver is on the right of Diablo and Fineas

---

[5] http://www.goats.com

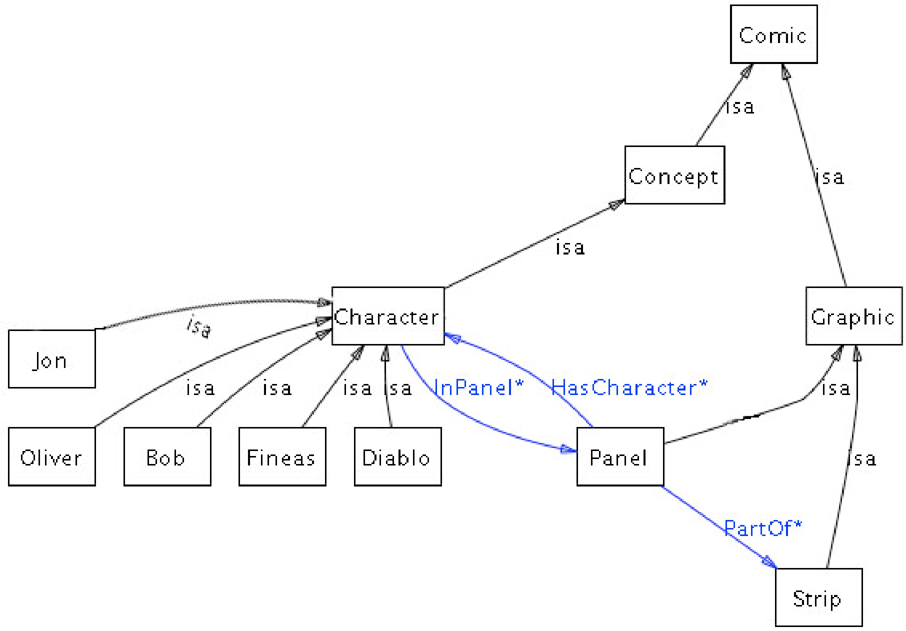

Figure 3. Visual graph of the image ontology generated using OntoViz (http://protege.cim3.net/cgi-bin/wiki.pl?OntoViz). Relationships — `isa`: `is-a`, `InPanel`, `HasCharacter`, and `PartOf`

Based on the tagged regions, we automatically calculated thecentroid coordinate of the region. The x-coordinate is used to determine the spatial location *left* and *right*. We defined *right* as the viewer's right (which is *stage left*). As depicted in Figure 2, the character, Diablo with the centroid coordinate (xc1, yc1), is to the left of Fineas with the centroid coordinate (xc2, yc2), because xc1 < xc2. Both characters, Diablo and Fineas are to the left of Oliver since xc1 and xc2 are less than xc3. The definition of spatial location *right* is the reverse of left. The region tags and their corresponding centroid coordinate are then incorporated into the ontology.

To support the spatial location query for a single character, we integrated the width information of the images into the ontology. When determining whether a character is located in the right or left of the image, we take the image width and divided it by two. Hence, we obtained the centre value of the image's width, cw, for the images. This is done during the retrieval process. Referring back to Figure 2, the comic character Oliver spatial location is in the right half of the panel. Our experiments related to spatial location only utilise the x-coordinate since most of the comic characters in the strips are located on the lower half of the panels. The upper half of the panels usually contain the speech balloon, which are not considered in our experiments, and hence ignored.

In building the ontology, we determine and classify the concepts related to the domain. Then, we recognise whether a concept has an `is-a` relationship with the superclass, for example, `character` is a subclass of `concept`. Figure 3 illustrates the ontology structure, and Figure 4 is a snippet of the OWL representation for the classes. Based on these two figures, the ontology structure is divided into two general classes `Concept` and `Graphic`. The class `Concept` has a subclass `Character` that further contains subclasses that represent the characters `Bob`, `Diablo`, `Fineas`, `Oliver` and `Jon`.

The `Graphic` subclasses are `Strip` and `Panel`. This ontology structure allows the graphic elements of the image collection to be separated from the semantic concepts. Classes in an ontology are connected using relationships. The class `Character` has an `InPanel` relationship with the class `Panel`, with the inverse relationship `HasCharacter`. The class `Panel` has a relationship of `Partof` with the class `Strip`. All other concepts are connected with an `is-a` relationship.

5. VOQL Specification

VOQL is a visual query language for querying image ontologies that are represented in both OWL and RDF format. It maps the visual queries to the SPARQL ontology query language. The advantages of using VOQL alongside VOQI are the ease of the ontology querying process and result interpretation. The user is able to formulate a query without any knowledge of the

```xml
<?xml version="1.0"?>
<rdf:RDF
   xmlns="http://dayang.cs.rmit.edu.au/~dayang/ComicOntology#"
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xmlns:owl="http://www.w3.org/2002/07/owl#"
   xmlns:p1="http://dayang.cs.rmit.edu.au/~dayang/ComicOntology.owl#"
   xmlns:p2="http://www.owl-ontologies.com/assert.owl#"
   xml:base="http://dayang.cs.rmit.edu.au/~dayang/ComicOntology">
<owl:Ontology rdf:about="The Goats Comic Ontology"/>
<owl:Class rdf:ID="Concept">
   <rdfs:subClassOf><owl:Class rdf:ID="Comic"/></rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Fineas">
   <rdfs:subClassOf><owl:Class rdf:ID="Character"/></rdfs:subClassOf>
</owl:Class>
   <owl:Class rdf:ID="Bob">
   <rdfs:subClassOf><owl:Class rdf:about="#Character"/></rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Jon">
   <rdfs:subClassOf><owl:Class rdf:about="#Character"/></rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Diablo">
   <rdfs:subClassOf><owl:Class rdf:about="#Character"/></rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Oliver">
   <rdfs:subClassOf><owl:Class rdf:about="#Character"/></rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Graphic"><rdfs:subClassOf rdf:resource="#Comic"/></owl:Class>
<owl:Class rdf:ID="Strip"><rdfs:subClassOf rdf:resource="#Graphic"/></owl:Class>
<owl:Class rdf:ID="Panel"><rdfs:subClassOf rdf:resource="#Graphic"/></owl:Class>
<owl:Class rdf:about="#Character"><rdfs:subClassOf rdf:resource="#Concept"/></owl:Class>
...
</rdf:RDF>
```

Figure 4. A snippet of the ontology class and subclasses in OWL format

SPARQL syntax via VOQI, which is then interpreted by VOQL. This prevents syntactic errors and a user would be able to focus on the information need, rather than on the formulation of query. Upon issuing a query through VOQI, VOQL converts the query into SPARQL and returns a set of answers that are presented visually, instead of a list containing the ontology elements. The Extended Backus-Naur Form (EBNF) notations are used to define the VOQL formal grammar. The definitions for the notations presented in Figure 5 according to Bray et al. (2006) are:

- ‘string’ matches a literal string that is given inside the single quotes.
- J|K, matches J or K. Where J and K represent simple expressions.

When a query is submitted to the retrieval component in Figure 1, VOQL interprets it into SPARQL. The SPARQL ontology query language is used to retrieve a set of answers that satisfies the query expressions. We used the ARQ Jena[6] query engine to process the SPARQL query and find all possible answers that match. A set of matching images are then displayed to the user.

```
[1] Query ::= Subquery | Subquery Operator Subquery
[2] Subquery ::= ComicCharacter | '(' Query ')'
[3] ComicCharacter ::= 'Bob' | 'Diablo' | 'Fineas' | 'Oliver' | 'Jon'
[4] Operator ::= 'AND' | 'OR' | 'NOT WITH' | 'RIGHT OF' | 'LEFT OF'
```

Figure 5. The VOQL formal grammar in EBNF

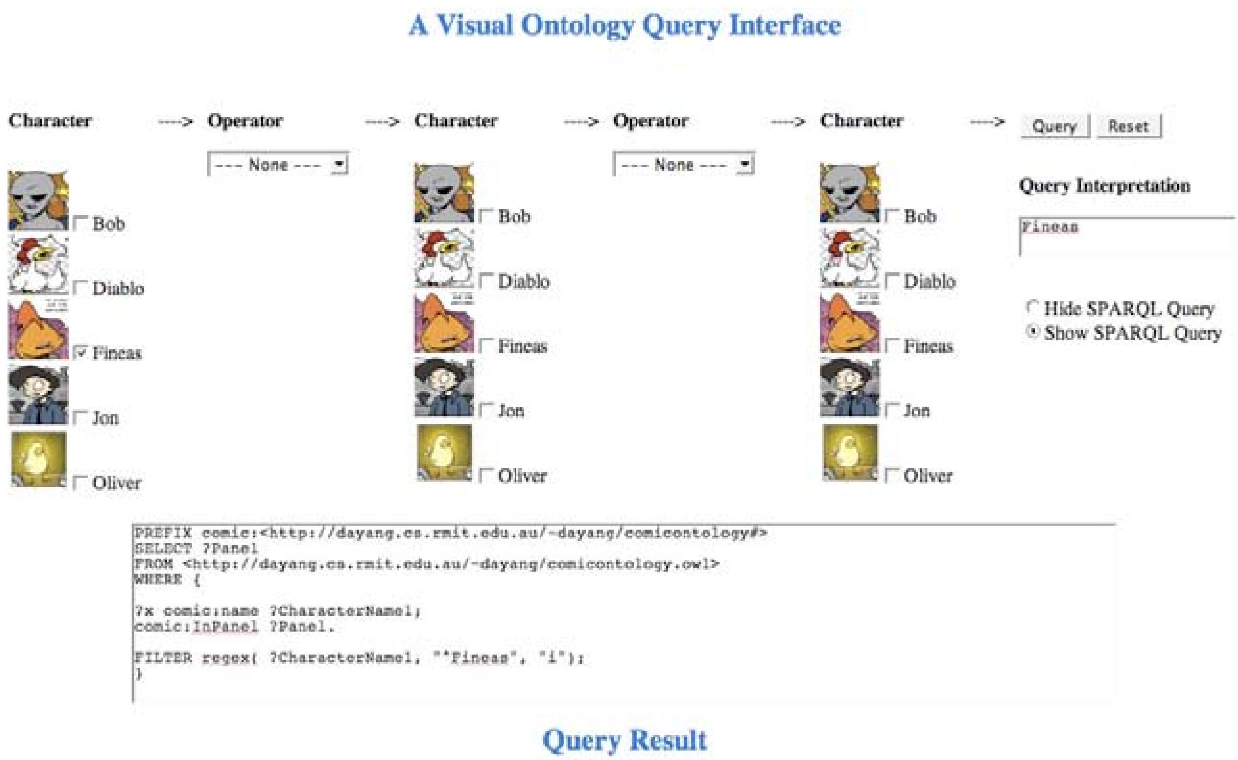

Figure 6. A sample query "Fineas" using VOQI. The answers depicted are only a small subset of the retrieved answers

6. CBIR Using VOQL

The Visual Ontology Query Interface (VOQI) is depicted Figure 6. This simple query interface supports up to 64 query combinations. Since the ontology was built to accommodate the knowledge of the regions in the image collection, the formulated queries can be seen as querying for the regions that represent the comic characters. These regions are represented by character icons. To handle the queries, we used VOQL, which is a formal text interpretation of the query expressed visually using VOQI. VOQL is more powerful than VOQI as it can represent every query generated by using VOQI, but is not limited to only three characters.

There are two types of queries. One is the character query, to find the desired characters in an image. The other is the spatial location query, which allows the user to pose a query that retrieves characters, based on their spatial location. To combine the characters, we use several combination operators:

- AND, to perform a logical conjunction on two expressions. We use this operator to retrieve images containing more than one character.
- OR, to perform a logical disjunction on two expressions. We use the exclusive OR operator (XOR) to retrieve images containing either one of the characters specified in the query.

[6] http://jena.sourceforge.net/ARQ

- NOT WITH, to perform logical negation on an expression. We use this operator to retrieve images containing other characters except the specified character.
- RIGHT OF, to perform the "right" positional expression; and
- LEFT OF, to perform the "left" positional expression.

When a query contains multiple combination operators, the system evaluates and resolves the query from left to right. Parentheses are used to force the first part of the query expression to be evaluated before other parts of the query. Operations within parentheses are always performed before those outside. In the case of a spatial location query, we evaluated the combination from left to right.

In this section, we present the result of single character, multiple characters and spatial location retrieval using VOQL. We used the standard recall and precision metrics. There are several points that need to be noted when evaluating the retrieval effectiveness of using an ontology. First, we retrieve all possible matching answers and secondly, these answers are not ranked in any order of similarity.

A sample of a single-subject character query "Fineas" and its answers is depicted in Figure 6. The SPARQL query interpreted using VOQL is:

```
PREFIX comic:
<http://dayang.cs.rmit.edu.au/˜dayang/ComicOntology#>
SELECT ?Panel
FROM <http://dayang.cs.rmit.edu.au/˜dayang/ComicOntology.owl>
WHERE{
   ?char comic:name ?CharacterName1;
   comic:InPanel ?Panel .
FILTER regex(?CharacterName1, "^Fineas", "i")) }
```

The SELECT query expression returns all the answers that are bound to the variable `?Panel`. The `WHERE` clause includes a set of graph pattern, `FILTER` and `regex` statement to match against. In this example, we have a set of two graph patterns:

- `?char comic:name ?CharacterName1`; is used to retrieve the name of the character; and
- `comic:InPanel ?Panel`. is used to retrieve the corresponding panels.

The FILTER term constraint is used to limit the return answers. Accordingly, `FILTER regex(?CharacterName1, "^Fineas", "i")` is used to retrieve panels that contain only the character `Fineas`, specified using the regular expression, which is case insensitive `("i")`. The SPARQL regular expression syntax follows the regular expression for XQuery 1.0 and XPath 2.0.[7]

Character	Recall	Precision
Bob	1.000	0.138
Diablo	0.900	0.704
Fineas	0.940	0.799
Jon	0.900	0.373
Oliver	0.966	0.350

Table 1. Retrieval Effectiveness of single character query

The retrieval effectiveness of retrieving a single character using the ontology from the image collection is shown in Table 1. We observe that the query for the character Bob has 100% recall, indicating that all the relevant images have been retrieved. On the other hand, the low precision value is due to large number of images retrieved as the answers. The second highest recall is for the character Oliver (96.6%), followed by the character Fineas (94%). Recall for both of the characters Diablo and Jon is the same (90%). The character Fineas had the highest precision (79.9%), which implies that most of the images retrieved contain Fineas.

[7] http://www.w3.org/TR/xpath-functions/#regex-syntax

Figure 7. A sample query answers for “Fineas LEFT OF Oliver” using VOQI.
The answers depicted are only a small subset of the retrieved answers

The interpretation of a multiple-subject character query (using AND, OR and NOT WITH operators) also uses the WHERE clause with several sets of graph patterns (FILTER and regex statements that were used in the singlesubject character query). The number of the graph patterns depend on the number of characters in the query. For example, a multiple-subject character query using the AND operator consists of the FILTER and regex statements of the single-subject character query repeated for each character in the query. A multiple-subject character query using the OR operator involves a UNION query condition. A multiple-subject character query using the NOT WITH operator, for example, “Fineas NOT WITH Diablo”, the first FILTER and regex statement is the same as the singlesubject character query and the second FILTER and regex statement include a negative regular expression denoted by a pair square bracket ([ ]). Therefore, the second FILTER and regex statement is written as FILTER regex(?CharacterName1, “[^Diablo]”, “i”).

A small set of retrieved images of a spatial location query for multiple characters is depicted in Figure 7. The query “Fineas LEFT OF Oliver” was interpreted using VOQL as:

```
PREFIX comic:
<http://dayang.cs.rmit.edu.au/~dayang/ ComicOntology#>
SELECT ?Panel
FROM <http://dayang.cs.rmit.edu.au/~dayang/>
WHERE {
  ?char1 comic:CentroidX ?CharacterCx1;
  comic:name ?CharacterName1;
  comic:InPanel ?Panel.
  ?char2 comic:CentroidX ?CharacterCx2;
  comic:name ?CharacterName2;
  comic:InPanel ?Panel.
FILTER regex( ?CharacterName1, “^Fineas”, “i”)
FILTER regex( ?CharacterName2, “^Oliver”, “i”)
FILTER (?CharacterCx1 < ?CharacterCx2) }
```

The WHERE clause includes two sets of graph pattern and three FILTER statements (two FILTER statements that include regex expression and one FILTER statement without any regular expression). The statements FILTER regex(?CharacterName1, “^Fineas”, “i”) and FILTER regex(?CharacterName2, “^Oliver”, “i”) specify that we want to retrieve images that contain the characters Fineas and Oliver. Their x-centroid coordinates were compared using FILTER (?CharacterCx1 < ?CharacterCx2). The SPARQL query for the RIGHT OF operator follows the same syntax as the query for the LEFT OF operator, except for the last FILTER and regex statement.

Instead of using <, the RIGHT OF operator uses > to compare the x-centroid coordinates. A spatial-location query for a single character was interpreted differently from the spatial-location query for multiple characters. Instead of comparing two x-centroid coordinates, the x-centroid coordinate of the character is compared with the centre panel width value. This is done during the retrieval process using the `FILTER (?CharacterCx1 < ?PanelWidth/2)` statement.

7. Conclusion and Future Work

In this paper we have presented a framework that integrates CBIR and ontology. We have described an ontology engineered using CBIR techniques to recognise objects that are in the images. The ontology is represented using OWL, a portable and machine independent ontology language. We have shown that synergy between ontology and image annotations is possible, and this method can reduce the gap between image features and high-level semantics, by providing the relationships between objects in the image. VOQI was integrated as the front-end for the user to query the ontology. This interface facilitates the ontology querying process, where there is no need for the user to master the SPARQL query language. Behind the interface is a visual query language, VOQL, that interprets the queries and represent them in SPARQL. VOQI and VOQL can also be promoted as the Web 2.0 applications, since they are built based on the semantic web technologies. Using robust object recognition techniques, the proposed ontology structure can be generalised to describe the objects and their locations for other image domains. Future directions for this work include: enhancing VOQI and refining VOQL semantic interpretation.

Acknowledgement

We acknowledge Universiti Malaysia Sarawak and RMIT University for the research support and facilities. Werecognise Assoc. Prof. James A. Thom and Dr S.M.M. Tahaghoghi for their contribution in this work. We also thankJonathan Rosenberg, author of the Goats Comic for the permission to publish the comic images.

References

[1] Awang Iskandar, D. N. F. (2009). Visual ontology query language, *In: Proceedings of the First International Conference on Networked Digital Technologies* (NDT 2009). Ostrava, Czech Republic.

[2] Awang Iskandar, D. N. F., Thom, J. A., Tahaghoghi, S. M. M. (2008a). Content-based image retrieval using image regions as query examples, *In:* Fekete, A., Lin, X. eds, 'Nineteenth Australasian Database Conference (ADC 2008)', *V. 75 of Conferences in Research and Practice in Information Technology*, Wollongong, Australia, p.38–46.

[3] Awang Iskandar, D. N. F., Thom, J. A., Tahaghoghi, S. M. M. (2008b). Query interface for content-based image retrieval ontology, *In: Proceedings of the 21st International Symposium on Human Factors in Telecommunication* (HFT 2008). Kuala Lumpur, Malaysia, p. 335– 342.

[4] Barnard, K., Forsyth, D. A. (2001). Learning the semantics of words and pictures, *In:* Werner, B. ed., *In: Proceedings of the Eighth International Conference on Computer Vision* (ICCV'01), V. 2, Vancouver, BC, Canada, p. 408–415.

[5] Bray, T., Paoli, J., Sperberg-Queen, C. M., Maler, E., Yergeau, F. (2006). Extensible Markup Language (XML) 1.0 (Fourth Edition), W3C Recommendation 16 August 2006, edited in place 29 September. http://www.w3.org/TR/2006/REC-xml-20060816

[6] Brewster, C., O'Hara, K., Fuller, S.,Wilks, Y., Franconi, E., Musen, M. A., Ellman, J., Shum, S. B. (2004). Knowledge representation with ontologies: The present and future, *IEEE Intelligent Systems,* p.72–81.

[7] EscÓrcio, A. L. N., Cardosa, J. (2007). Editing tools for ontology creation, Semantic Web Services: Theory, Tools and Applications, Information science reference, chapter 4, p. 71–95.

[8] Gandon, F. (2002). Ontology engineering: a survey and a return on experience, Technical Report 4396, INRIA.

[9] Hare, J. S., Lewis, P. H., Enser, P. G. B., Sandom, C. J. (2006). Mind the gap: Another look at the problem of the semantic gap in image retrieval, *In:* Chang, E. Y., Hanjalic, A., Sebe, N., ed., *Proceedings of Multimedia Content Analysis, Management and Retrieval*, V. 6073.

[10] Hare, J. S., Sinclair, P. A. S., Lewis, P. H., Martinez, K., Enser, P. G., Sandom, C. J. (2006). Bridging the semantic gap in multimedia information retrieval: Top-down and bottom-up approaches, *In*: Bouquet, P., Brunelli, R., Chanod, J.-P., Nieder´ee, C., Stoermer, H. eds, *Mastering the Gap: From Information Extraction to Semantic Representation, Proceedings of the Third European Semantic Web Conference*, Budva, Montenegro.

[11] Harit, G., Chaudhury, S., Paranjpe, J. (2005). Ontology guided access to document images, *In:* Werner, B. ed., *Proceedings of the Eighth International Conference on Document Analysis and Recognition* (ICDAR '05), Seoul, Korea, p.292–296.

[12] Hollink, L., Schreiber, G., Wielemaker, J., Wielinga, B. (2003). Semantic annotation of image collections, *In:* Handschuh, S., Koivunen, M.-R., Dieng, R., Staab, S. eds, *In:* Proceedings of the of the K-CAP 2003 *Workshop on Knowledge Markup and Semantic Annotation* (Semannot '2003). Sanibel, Florida, USA.

[13] Hyvönen, E., Saarela, S., Viljanen, K. (2003). Ontogator: Combining view- and ontology-based search with semantic browsing, *In*: *Proceedings of XML Finland* 2003, Kuopio, Finland.

[14] Jelmini, C., Marchand-Maillet, S. (2003). Deva: an extensible ontology-based annotation model for visual document collections, *In:* Santini, S., Schettini, R. eds, SPIE, V. 5018.

[15] Jiang, S., Huang, T., Gao, W. (2004). An ontology-based approach to retrieve digitized art images, in Zhong, N., Tirri, H., Yao, Y., Zhou, L., Liu, J., Cercone, N. eds, WI'04: *Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence*, Beijing, China, p.131–137.

[16] Liu, S., Chia, L.-T., Chan, S. (2004). Ontology for nature scene image retrieval, *In:* Meersman, R., Tari, Z. eds, *On the Move to Meaningful Internet Systems: Proceedings of the OTM Confederated International Conferences CoopIS, DOA, and ODBASE 2004*, V. 3291 of Lecture Notes in Computer Science, Agia Napa, Cyprus, p.1050–1061.

[17] Schreiber, A. T. G., Dubbeldam, B., Wielemaker, J., Wielinga, B. (2001). Ontology-based photo annotation, IEEE Intelligent Systems 16 (3) 66–74.

[18] Tamura, H., Mori, S. (1977). A data management system for manipulating large images, *In: Proceedings of Workshop on Picture Data Description and Management*, Chicago, Illinois, USA, p.45–54.

[19] Tamura, H., Yokoya, N. (1984). Image database systems: A survey, Pattern Recognition 17 (1) 29–43.

[20] Town, C. P. (2006). Ontological inference for image and video analysis, Machine Vision and Applications 17 (2) 94– 115.

[21] Troncy, R., van Ossenbruggen, J., Pan, J. Z., Stamou, G. (2007). Image annotation on the semantic web, W3C incubator group report'. URL: http://www.w3.org/2005/Incubator/mmsem/ XGR-image-annotation/.

[22] Wang, H., Liu, S., Chia, L.-T. (2006). Does ontology help in image retrieval? A comparison between keyword, text ontology and multi-modality ontology approaches, *In:* Zimmermann, R. ed., *MULTIMEDIA '06: Proceedings of the 14th annual ACM International Conference on Multimedia*, Santa Barbara, CA, USA, p.109–112.

Author biographies

Dayang NurFatimah Awang Iskandar is a lecturer at the Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak. She has a bachelor degree in information technology, a master degree in multimedia computing, and a PhD in computer science. Her research is currently focused on minimising the gap between image features and high-level semantics in content-based image retrieval systems. She has worked on techniques for querying an image collection including: fusing text accompanying the images with visual features, automatic region tagging, and using ontology to enrich the semantic meaning of tagged image regions leading to the bridging of semantic gap in content-based image retrieval.